Routing in IP Netzen

Felix von Leitner
Chaos Computer Club Berlin
felix@ccc.de

Chaos Communication Congress 1998

Zusammenfassung

Routing beantwortet die Frage: *Wieso kommt mein Paket nicht an?*

Wie spricht man das eigentlich aus?

[rauting] und *[ruting]* sind beide richtig. Amerikaner benutzen gerne letzteres, Europäer gerne ersteres.

Router und Routing-Protokolle

Eine **Route** ist eine Information, die einem Gerät sagt, wie es ein Paket in ein bestimmtes Ziel-Subnetz senden kann. Eine Route heißt **asymmetrisch**, wenn die Rückroute einen anderen Pfad benutzt.

Ein **Router** ist ein Gerät, das zwei Subnetze verbindet. Geräte in beiden Subnetzen geben ihre Pakete für das jeweils andere Subnetz beim Router ab.

Ein **Routing-Protokoll** ist ein Mechanismus für das dynamische Entdecken der Pfade für Datenpakete durch das Internet.

Ein **Hop** ist ein Gerät auf dem Pfad zwischen zwei Geräten.

IPv4-Adressen

- Eine IPv4-Adresse ist ein 32-bit Integer
- Wird gewöhnlich in der Form **192.168.17.23** geschrieben (**192** ist das höchstwertigste Byte)
- Im IP-Header steht sie in der *network order* (d.h. big-endian)
- IPv4-Adressen sind im Internet eindeutig

IPv6-Adressen

- Eine IPv6-Adresse ist ein 128-bit Integer
- Teil der IPv6-Adresse ist die (eindeutige) MAC-Adresse

Alle Betrachtungen in diesem Vortrag sind bezüglich IPv4!

Was ist eine Netzmaske?

Die Netzmaske wird benutzt, um die Netzadresse zu berechnen. Dafür wird eine IP-Nummer mit der Netzmaske logisch **AND**-verknüpft.

Beispiel:

IP-Nummer	192	.	168	.	17	.	23
Netzmaske	255	.	255	.	0	.	0
Netzadresse	192	.	168	.	0	.	0

Tabelle 1: Netzadresse berechnen

Wozu braucht man die Netzadresse?

Zwei IP-Nummern sind im gleichen Subnetz, wenn sie bezüglich der gleichen Netzmaske die gleiche Netzadresse haben.

Per Konvention haben Netzmasken die Binär-Form `1*0*`, d.h. die Einsen sind alle linksbündig, aber technisch ist das nicht notwendig. Wenn Netzmasken in dieser Form sind, schreibt man sie auch in der Form **10.0.0.0/8** (für die Netzmaske 255.0.0.0).

Eine Netzadresse und eine Netzmaske zusammen bestimmen ein **Subnetz** eindeutig.

Ein Subnetz und eine IP-Adresse zusammen ergeben eine **Route**, wobei die IP-Adresse die des **Gateways** ist. Das Gateway ist ein Gerät, das einen **Hop** näher am Ziel ist. Lokale Netze (d.h. solche, bei denen der Rechner selbst Mitglied ist) haben kein Gateway.

Eine Route mit einer Netzmaske von **0** wird **default** Route genannt, weil sie bei allen Ziel-Adressen zutrifft.

Was ist denn ein Subnetz?

Es gibt nur 5 Klassen von Netzen, die vom InterNIC vergeben werden:

Klasse	IP-Bereich	Netzmaske
Class A	0-127.*	255.0.0.0
Class B	128-191.*	255.255.0.0
Class C	192-223.*	255.255.255.0
Class D (multicast)	224-239.*	255.0.0.0
Class E (reserviert)	240-255.*	255.0.0.0

Tabelle 2: Netzklassen

Man kann natürlich andere Netzmasken benutzen. Die sich ergebenden Netze heißen (aus historischen Gründen) Subnetze. Heute ist diese Trennung unwichtig und man nennt beides einfach Netze.

Cisco und Subnetze

Manche Hersteller optimieren die (sequentielle) Suche in großen Routing-Tabellen, indem sie nur die Netzadressen vergleichen, was bei **10.0.0.0/8** und **10.0.0.0/24** u.ä. nicht funktioniert.

Bei Cisco läßt sich dieser Bug mit **ip subnet-zero** korrigieren.

Was ist eine Routing-Tabelle?

Jedes IP-Gerät hat eine Liste von Routen. Wenn ein Paket verschickt werden soll, wird für alle Routen geprüft, ob die Ziel-Adresse in dem entsprechenden Netz liegt. Wenn ja, wird es an die IP des Gateways der Route geschickt.

Damit nicht immer die Default-Route angewandt wird, sind Routen noch mit einer **metric** gewichtet. Eine höhere Metrik steht für eine teurere Route. Die Default-Route hat deshalb gewöhnlich Metriken ungleich Null.

Das kann aber auch benutzt werden, um Backup-Routes über langsamere Geräte zu definieren, die nur benutzt werden, wenn die Hauptroute ausfällt.

Zusätzlich steht bei jeder Route noch das Netzwerk-Interface dabei, über das das Paket mit der IP-Nummer des entsprechenden Gateways gesendet werden soll. Diese wird normalerweise anhand früherer Routen und der Gateway-IP bestimmt.

Des Pudels Kern

Dieser Vortrag behandelt die Frage, wo der Inhalt der Routing-Tabelle herkommt, bzw. wie man Geräte überredet, anders als in der Routing-Tabelle beschrieben zu routen.

Tatsächlich interessiert den Angreifer gewöhnlich nur die Rück-Route, d.h. man möchte gewöhnlich erreichen, daß Pakete bei einem Router vorbeigeroutet werden, auf den man privilegierten Zugriff hat.

Statisches Routing

Die einfachste Methode für das Bevölkern der Routing-Tabelle sind statische Routen, d.h. von Hand eingetragene. Die Routen für die lokalen Netzwerk-Interfaces müssen immer statisch definiert werden. Normalerweise können Routing-Protokolle keine statischen Routen ändern.

Dynamisches Routing

Dynamisches Routing heißt, daß die Routing-Tabelle über das Netzwerk modifizierbar ist. Verbreitete Protokolle dafür sind

- ICMP Redirect
- RIP (Routing Information Protocol)
- OSPF (Open Shortest Path First)
- BGP (Border Gateway Protocol)
- IGRP (Interior Gateway Routing Protocol)
- EIGRP (Enhanced IGRP)

Was kann man alles falsch machen?

Das offensichtliche Routing-Protokoll wäre, daß alle Rechner periodisch ihre Routen broadcasten und die bekannten Routen mit einem Timeout belegen.

- Yoyo-Effekt
- Routen verschwinden nicht
- Falschrouten propagieren sich

Yoyo-Effekt

1. - Rechner A hat eine Leitung zu Rechner C
   - Rechner A hat eine Route über diese Leitung, Metrik 1
   - Rechner B kennt die Route mit Gateway A und Metrik 2
2. Die Leitung bricht zusammen, Rechner A löscht die Route
3. Rechner B nennt Rechner A die Route
4. Rechner A kennt jetzt eine Route über die tote Leitung mit Rechner B als Gateway und Metrik 3

Routen verschwinden nicht

Wenn ein Rechner seine Route löscht, bekommt er sie umgehend von einem Nachbarn mit dem nächsten Broadcast wieder mitgeteilt.

- Man könnte periodisch die Metrik verschlechtern bei den alten Routen. Dann würden Schrott-Routen irgendwann schlechter als die Default-Route.
- Man könnte verbieten, daß Routen an den Rechner zurückgemeldet werden, von dem sie kommen. Damit verhindert man aber keine größeren Schleifen (z.B. drei Rechner reichen eine Route im Kreis)
- Man könnte negative Routing-Meldungen einführen, d.h. eine „bitte löschen“ Meldung. Die müßte auch nach einer Weile gelöscht werden, sonst akkumulieren sich Müll-Broadcasts.

Warum überhaupt dynamisch routen?

- Automatisch auf Backup-Pfad umschalten
- Besten Pfad aus mehreren Alternativen wählen
- Last auf mehrere Pfade verteilen (nur dedizierte Router)
- Weniger administrativer Aufwand

Source Routing

Es gibt im IP-Header die Möglichkeit, ein paar Hops der Rückroute zu spezifizieren, und dem Zielrechner zu sagen, daß die Rückroute doch bitte so aussehen soll. Weil dafür praktisch keine legitimen Gründe für die Anwendung existieren, lassen praktisch alle Firewalls und die meisten Router solche Pakete auf den Boden fallen.

Mit Source-Routing kann man mit **netcat** herumspielen.

ICMP Redirect

ICMP ist das IP-Protokoll für Kontrollnachrichten. Es wird für *ping* und für *port unreachable*, *network unreachable*, *host unreachable* und eben *redirect* benutzt.

Mit der letzten Nachricht kann der Gateway einem Host sagen, daß er (im lokalen Netz) doch nicht zuständig ist und für das Ziel doch lieber dieser andere Gateway zuständig wäre.

Leider akzeptieren manche Implementationen auch ICMP redirect Pakete von anderen Rechnern als dem zuständigen Gateway (Router inzwischen nicht mehr). ICMP redirect verstehen auch Nicht-Router, Router unterhalten sich anders.

ICMP redirect Pakete kann man von Hand basteln, mit **spak** und vor kurzem wurde auch ein Perl-Modul für den Hobby-Paketbastler angekündigt.

Klassen von Routing-Protokollen

Man unterscheidet zwischen *Interior* und *Exterior* Routing-Protokollen. Erstere werden innerhalb von *autonomous systems* benutzt, letztere verwalten den Verkehr zwischen diesen.

Autonome Systeme sind z.B. MILNET und NSFNET, d.h. normalen Administratoren begegnen nur Protokolle vom Typ *Interior*. Autonome Systeme werden anhand einer 16-bit Zahl unterschieden, die man offiziell beim InterNIC beantragen muß.

Das meistgenutzte *interior* Protokoll ist RIP, man kann aber auch OSPF benutzen.

Routing Information Protocol

- RIP wurde beim XEROX PARC entwickelt und 1981 für IP formal definiert (RFC 2453; auch 1058, 1388, 1723)
- Es definiert ein Format, mit dem man sagen kann, daß man eine Route zu Rechner XY mit Metrik Z kennt, oder fragen kann, ob jemand eine Route zu Rechner XY kennt
- Nachrichten gehen nur an Nachbarn im LAN
- Die Metrik einer Route wird beim Import inkrementiert, d.h. sie entspricht dem Hop-Count
- Das nennt man „distance vector“
- Nachrichten sind trivial spoofbar
- RIP 2 definiert ein Paßwort-Feld, das natürlich auch snoopbar ist
- RIP definiert ein paar triviale Heuristiken, mit denen Loops verhindert werden sollen
- Der BSD **routed** benutzt RIP

RIP Heuristiken

- Hops sind limitiert auf 15, d.h. unbenutzbar für sehr große Netze und das ganze Internet
- Hold-Down (Routen werden nicht gelöscht, sondern als gelöscht markiert und eine Weile eingefroren)
- Split Horizon (Routen nicht an deren Gateway propagieren)
- Poison Reverse Updates (Beim Nachbarn Routen mit mehr Hops löschen)

Open Shortest Path First

- Von IETF entwickelt, weil RIP nicht skalierbar genug war (RFC 2328; auch 1131, 1247, 1583, 2178)
- Schickt nicht nur eigene Routen und nicht nur an Nachbarn im LAN
- Jeder Router akkumuliert einen Graphen aus Routen
- In diesem Graphen wird die richtige Route mit Dijkstras Algorithmus gesucht
- Ein autonomes System kann aus mehreren „areas“ bestehen
- Router in mehreren Areas heißen „area border routers“ und halten einen Graphen pro Area
- Nachbarn werden mit dem *OSPF Hello protocol* gefunden, eine Art *ping*, das auch als keepalive benutzt wird
- OSPF unterstützt *type of service* Routing mit *delay*, *throughput* und *reliability* als mögliche Anforderungen
- *Equal cost multipath routing* (nur dedizierte Router)

Interior Gateway Routing Protocol

- Cisco-proprietäres „distance vector“ Protokoll
- Metrik-Vektor: *delay*, *bandwidth*, *reliability* und *load*
- Kann *multipath routing* (channel bundling)
- Push-Protokoll, ausgefallene Router werden an fehlenden Updates erkannt
- Zusätzlich zu den RIP-Heuristiken gibt es verschiedene Timer

Enhanced IGRP

- Merkt sich die Routing-Tabellen aller Nachbarn
- Subnetzmasken variabler Länge
- Keine periodischen Updates, partielle on-demand Updates
- Periodische Hello-Pakete für Discovery
- *Reliable Transport Protocol* (RTP) sorgt für garantierte in-order Auslieferung der IGRP-Pakete

Exterior Protocols

Diese Protokolle werden von wenigen zentralen Routern benutzt, die von kompetenten Leuten gepflegt werden.

Spoofing ist gewöhnlich zwecklos, und man kann damit auch nur bei den Übergängen zwischen autonomen Systemen spoofen, d.h. es wäre eh nicht sehr hilfreich.

Nicht-dedizierte Router benutzen gewöhnlich die **gated**-Implementation. Neuerdings wird auch Multicasting benutzt für Exterior Protocols.

Exterior Gateway Protocol

- Erreichbarkeit, nicht Routing, von 1984 (RFC 0904; auch 0827)
- Periodisches *Hello/I-Heard-You*
- Polling
- Definiert über einen endlichen Automaten
- Nachrichten nur zwischen je zwei direkten Nachbarn
- EGP gibt auch Routen anderer Leute weiter
- Router werden direkte Nachbarn mit einem 3-way handshake
- Direkte Nachbarn pollt man periodisch nach deren Routen

EGP definiert nicht, wie man auf die Nachrichten reagieren soll, oder daß man die Routing-Tabelle überhaupt anfassen soll.

Border Gateway Protocol

- Ein „Exterior Gateway Protocol“ von 1989 (RFC 1771; auch 1105, 1163, 1267, 1654)
- Routing zwischen autonomen Systemen
- Nachfolger des RFC0904-EGP
- Kann Routing-Loops erkennen
- Keine periodischen Table-Broadcasts
- BSP merkt sich für alle Peers deren aktuelle Routing-Tabelle
- Inkrementelle Update-Nachrichten melden jeweils den optimalen Pfad, nicht alle Pfade
- Pakete beinhalten 16 Bytes „Marker“ für Authentisierung und ein Feld, mit dem man den Authentisierungs-Algorithmus nennen kann, aber keine Algorithmen sind definiert
- Für die Verbindungen wird TCP benutzt

Welches Protokoll soll man denn benutzen?

In fast allen Fällen ist RIP vollständig ausreichend, und nicht leichter spoofbar als OSPF. Man kann mit **gated** und **routed** RIP fahren, und wenn man die Wahl hat, sollte man **gated** wählen.

Wenn man schon RIP benutzt, sollte man auf jeden Fall statische ARP-Einträge für die bekannten Router benutzen und IPsec o.ä. benutzen zwischen den Routern.

Spoofing mit Routen

- RIP-Routen sind trivial zu fälschen
- Sie propagieren sich dann aber auch brav
- ...und halten sich auch eine Weile!
- Der Fälscher ist nur anhand der Route nachvollziehbar, weil Routen-Spoofing gewöhnlich Traffic durch ein Netz zwingt, auf dem der Angreifer privilegierten Zugriff hat